Web Version

Read On!

Research Compilation Report

November 2005

Harcourt Achieve
Rigby • Saxon • Steck-Vaughn

This page intentionally left blank.

Read On!™ Research Compilation Report

Table of Contents

Introduction

In 2005, Steck-Vaughn published *Read On!*, the latest revision of the *Learning 100 System* reading program. Although much of the content, all of the graphics and interface, and the back-end architecture were changed to bring to students and teachers an up-to-date and exciting program, the pedagogical model—which has a 40-year record of success in the remediation of struggling readers—in the newly renamed program remains unchanged. Because the instructional architecture of *Read On!* evolved from and is based on its predecessor, *Learning 100*, the instructional effectiveness of the *Learning 100* program can provide a solid foundation for evaluating the efficacy of *Read On!* This report compiles recent research findings from sites across the United States that have implemented the *Learning 100* program. In each case, *Learning 100* was shown to be effective at impacting growth in student reading achievement for each variable that was investigated, for various populations of students, and for different lengths of implementation time. Overall, the pedagogical model of instruction that is the basis of *Read On!* and its predecessor, *Learning 100*, is an effective instructional model that has been shown to improve the reading ability of struggling readers.

Read On! was also evaluated to examine the alignment between the *Read On!* program and the key instructional and infrastructional recommendations for successful middle and high school reading instruction put forth by the Alliance for Excellent Education in their report, *Reading Next: A Vision for Action and Research in Middle and High School Literacy*. The *Reading Next* recommendations provided the framework around which the scientific research base of the *Read On!* program was established. This report, also available from Harcourt Achieve, provides an overview of how *Read On!* incorporates the *Reading Next* instructional and infrastructional recommendations that contribute to the reading success of struggling adolescent readers by demonstrating clearly and explicitly how *Read On!* incorporates these recommendations into its instructional model. Please visit www.harcourtachieve.com for this report.

Hawaii

A multi-group evaluation of the *Learning 100* program was conducted at a high school/intermediate school on the Oahu island of Hawaii in a small urban city on the fringe of a larger city. Students from seventh to tenth grade used the program throughout the 2003-2004 academic school year. At the time of the evaluation, there were over 700 students enrolled in the school, with approximately 51% of the total student population qualifying for free or reduced lunch. The dominant ethnicities in the school accounted for 74% of the student population. Thirty-six percent of these students were of Filipino decent, 12% Caucasian, 26% Hawaiian or part Hawaiian. The rest of the school population was comprised of several smaller ethnicities, including Japanese, Laotian, and Cambodian. The demographics of the school have been stable over the past six years, including the ESL population, which has held at about 7.1%.

Of the 700 students in the school, 284[1] were selected to participate in the *Learning 100* implementation after being identified as reading below grade level. In an effort to meet the needs of all of the struggling readers, there was a differential implementation of the *Learning 100* program based on sub-populations of the students across all grade levels. For most of the participants, *Learning 100* was implemented as a separate class in an extended-day setting. During the 2002-2003 school year, the district had implemented the *Learning 100* program as a pull-out program for regular education students who were reading below grade level. However, the administrators felt that they could service the needs of more students and achieve greater success with the program by creating a separate *Learning 100* class. The class was conducted twice a week for 45 minutes each session. A majority of the Regular Education students who were reading below grade level (n =165) used the *Learning 100* program as a stand-alone intervention and only attended the extra *Learning 100* class. However, for a subset of the Regular Education students who were reading below grade level (n = 42), *Learning 100* was provided during the regular school day as part of their reading curriculum. For these students, the *Learning 100* program was implemented for 50% of the classroom time and for the other 50% of the time the students received reading instruction via traditional methods. Participating students who were classified as Special Education (n = 50) or ELL (n = 27) participated in the *Learning 100* program in addition to their classroom instructor-led reading curriculum[2].

All students attending the school were administered the STAR Reading Assessment as a pre- and posttest measure to determine the reading achievement gains over the 2003-2004 school year. Results of this assessment were also used to determine the gains in achievement that were specifically due to the implementation of the *Learning 100* program. Collectively, the students who participated in the *Learning 100* program had an average initial score of 544 (*SD* = 168.83) on the STAR Reading Assessment. After using the *Learning 100* program for one year, the scores of these students rose to an average of 655 (*SD* = 216.57). This gain was highly significant at the traditional alpha level of .05[3] ($t(283) = 12.69$, $p < .001$; Cohen's *d* Effect Size of .57[4]). In terms of grade equivalent scores, students scored on average a 4.8 on the pretest and excelled to an average of 6.0 on the posttest. In general, expected

[1] Originally there were 285 students who participated in the *Learning 100* program; however, one of the students had scores on the STAR pre- and posttest that were considered to be extreme outliers in the population. This student's scores were removed from all analyses in order to provide a more representative analysis not skewed by the performance of one student.

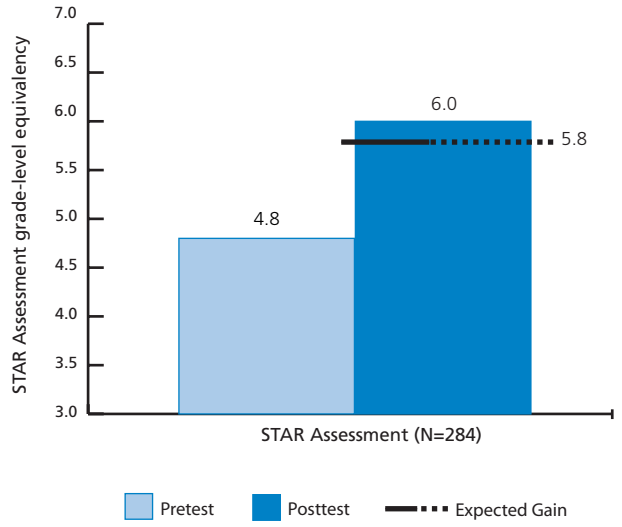
[2] These students were receiving instruction from the Direct Instruction and SRA programs, both of which are highly scripted reading intervention programs.

[3] An alpha level of .05 was used for all evaluations of statistical significance in this report unless otherwise noted.

[4] All effect sizes in this report were calculated using Cohen's *d* formula unless otherwise indicated. The results should be interpreted in terms of the following guidelines: .2 is equal to a small effect, .5 is a moderate effect, and .8 and above is a large effect. In practical terms, an effect size of .2 indicates that the average member of the group with the larger mean would score higher than 58% of the members in the group with the lower mean. For an effect size of .5, that individual would score higher than 69% of the lower performing group, and for an effect size of .8, members in the higher performing group would score higher than 79% of the members in the group with the lower mean score. It should be noted that according to Slavin (1986), a leader in educational research, an effect size of .25 is considered educationally significant.

grade equivalency gain over time on the STAR Reading Assessment without the use of a specialized program to improve reading achievement is 1.0. Students participating in the *Learning 100* program achieved slightly better scores than would be expected by gaining 1.2 grade levels (See Figure 1). Given that these students were struggling readers, a slightly better than normal gain in grade-level equivalency is an outstanding accomplishment. Only 26% of the participating students saw no gain in grade-level equivalency. Of the 74% of participants whose scores increased, 48% of them experienced gains of one full grade level or more, with 15%[5] obtaining on-grade-level scores.

Figure 1
Growth in grade-level equivalency for all students after using *Learning 100* for one year



In order to determine if there was a differential effect of the *Learning 100* program in relation to the grade level of the student, participants' scores on the STAR pre- and posttest were examined. There were 59 seventh-grade students, 71 eighth-graders, 60 ninth-graders, 47 tenth-graders, and 16 eleventh-grade students. Thirty-one of the students who participated in the *Learning 100* program were unclassified as to their grade level and not included in this analysis. Students' average gain on the STAR Assessment across grades was examined. Students' average scores on the STAR pre- and posttest assessment can be found in Table 1.

Table 1
STAR Assessment Pre- and Posttest Average Scale Score by Grade Level

Grade	N	Mean	Standard Deviation
STAR Pretest			
Seventh	59	489	152.39
Eighth	71	571	152.23
Ninth	60	582	129.65
Tenth	47	591	189.27
Eleventh	16	577	268.48
Total	253	558	167.51
STAR Posttest			
Seventh	59	617	206.86
Eighth	71	691	212.79
Ninth	60	701	169.43
Tenth	47	677	255.99
Eleventh	16	714	283.66
Total	253	675	216.86

Because it was expected that students would score differently on the pre- and posttests across grade level, the differential performance on the STAR Assessment was not conducted between grade levels. Instead, the change from pre- to posttest was calculated and averaged across grade levels (see Table 2).

Table 2
Average Change on the STAR from Pre- to Posttest by Grade Level

Grade	N	Mean	Standard Deviation
Change Score			
Seventh	59	128	151.16
Eighth	71	120	162.71
Ninth	60	119	138.08
Tenth	47	87	156.11
Eleventh	16	120	162.71
Total	253	117	151.29

[5] This proportion is based on a subset for which grade level information was provided and not the whole population.

A univariate ANOVA found that there was no difference in the amount of change from pre- to posttest among the grade levels, indicating that all students who participated in the *Learning 100* program progressed equally over time and that *Learning 100* did not have a differential effect in specific grade levels. However, as can be seen in Table 3, in terms of grade equivalency, students in all grade levels except tenth grade achieved over a 1.0 grade-level gain on average. Any gains over 1.0 are considered higher than would be expected. See Figure 2.

Table 3
Average Grade Equivalency Change in the STAR from Pre- to Posttest by Grade Level

Grade	N	Mean
	Change Score	
Seventh	59	1.3
Eighth	71	1.1
Ninth	60	1.1
Tenth	47	0.8
Eleventh	16	1.3
Total	253	1.2

Because the population of students participating in the *Learning 100* program implementation were all classified as struggling readers, this population included Regular Education students reading below grade level, as well as Special Education and ELL students also reading below grade level. Of the 284 students participating in the *Learning 100* program, there were 207 Regular Education students, 50 Special Education students, and 27 ELL students. Of the 207 Regular Education students, 165 of them were the students who used the *Learning 100* program as a standalone intervention; the other 42 participated in the *Learning 100* program in addition to focused curriculum instruction in reading. Table 4 displays the average scores on the STAR Assessment pre- and posttest for these groups.

A univariate ANOVA found that there was significant difference among the groups' pretest scores on the STAR Assessment ($F(3,280) = 56.89$, $p < .001$, partial eta squared effect size = .39[6]). Post-

[6] The effect size was calculated using the partial eta squared formula, which calculates the proportion of total variability that is attributable to the factor being analyzed. It ranges from 0.00 to 1.00 and should be interpreted like a percentage.

Figure 2
Actual versus expected grade-level gain on the STAR Assessment after using *Learning 100*

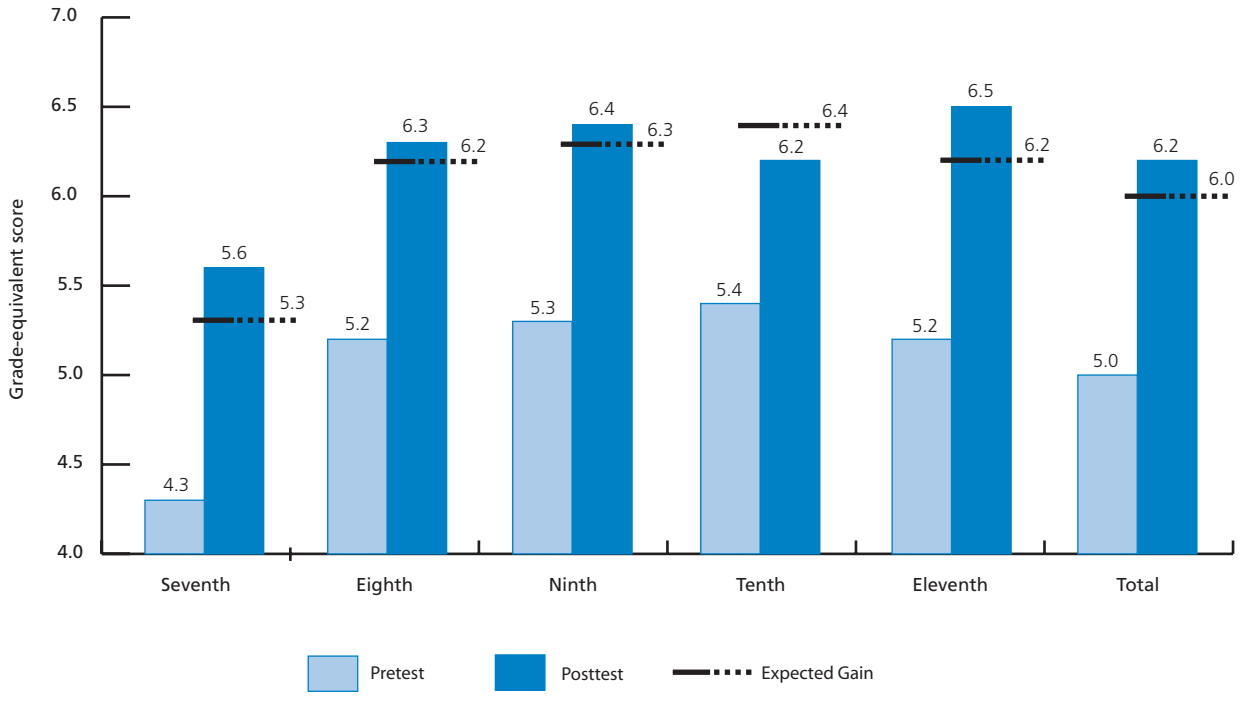


Table 4
STAR Pre- and Posttest Average Scale Score by Sub-Group Classification

Classification	Mean	Standard Deviation	N
STAR Pretest			
Regular Education	629	119.50	165
Regular Education Plus Classroom Instruction	475	185.98	42
Special Education	388	129.14	50
English Language Learner	416	129.26	27
Total	544	168.83	284
STAR Posttest			
Regular Education	744	194.53	165
Regular Education Plus Classroom Instruction	654	205.61	42
Special Education	449	122.81	50
English Language Learner	492	130.27	27
Total	655	216.57	284

hoc analysis using Gabriel's pairwise comparisons[7] found that both of the Regular Education groups had higher pretest scores than the Special Education group and the Regular Education group with the stand-alone *Learning 100* intervention had significantly higher pretest scores than both the ELL students and the Regular Education students with the additional classroom reading instruction (all differences significant at the alpha .05 level). However, given the analysis includes comparisons of intact groups, an ANCOVA[8] was not performed on the data. Instead, a univariate ANOVA was conducted to look at the difference between the groups on average gain scores between the pre- and posttest. On average, the Regular Education students in the stand-alone *Learning 100* condition had a gain of 115 (*SD* = 149.07) points, Regular Education students enrolled in the additional reading curriculum had an average gain of 179 (*SD* = 181.83) points, Special Education students gained on average 61 (*SD* = 107.18) points, and ELL students had an increase on average of 75 (*SD* = 100.63) points. The difference among these means was significant ($F(3,280) = 5.70$, $p = .001$; partial eta squared effect size = .06). This analysis indicates that the groups did achieve significant differential gains on the STAR Assessment posttest depending on group classification. A post-hoc analysis using Gabriel's pairwise comparisons revealed that the difference was generated from the Regular Education students who were provided with the additional classroom-based reading instruction. This group achieved significantly more growth than the other three groups (significant at the alpha .05 level).

Average grade-equivalent scores were obtained for each subgroup classification for the pre- and posttest STAR Assessment based on the groups' average scale score (see Table 5). As can be seen in Figure 3, both Regular Education groups achieved average gains on target or higher than the expected 1.0 grade-level increases. The Regular Education group receiving the additional classroom instruction obtained higher than expected gains increasing by almost two grade levels. The amount of expected increase for the Special Education and ELL groups cannot be determined, although it is likely that they would not be expected to increase a full grade level over the course of a year.

Given that the Regular Education students who received the additional classroom instruction in reading skills had larger gains from pre- to posttest

[7] Gabriel's pairwise comparison was chosen as the post-hoc statistic because the large variance in sample size across the grade levels.

[8] It is not recommended that an ANCOVA be performed on intact groups even though the pretest scores are significantly different from one another. The ANCOVA analysis essentially equates the pretest scores as an average individual between the groups and, in this case, in the actual population, such an average person would not likely exist, given the groups are for the most part mutually exclusive.

Table 5
STAR Pre- and Posttest Average Grade-Level Equivalency Scores by Sub-Group Classification

Classification	Mean	N
STAR Pretest		
Regular Education	5.7	165
Regular Education Plus Classroom Instruction	4.2	42
Special Education	3.3	50
English Language Learner	3.6	27
Total	4.8	284
STAR Posttest		
Regular Education	6.7	165
Regular Education Plus Classroom Instruction	6.0	42
Special Education	3.9	50
English Language Learner	4.3	27
Total	6.0	284

than other struggling readers, a more in-depth examination was undertaken to determine if the additional classroom instruction benefited students more than just *Learning 100* alone. Regular Education students who additionally had the classroom instruction gained more from pre- to posttest than the Regular Education students who only had *Learning 100* instruction. However, after pretest scores were examined, Regular Education students with the additional classroom instruction had a pretest score average of 475 (SD = 185.98), while Regular Education students with *Learning 100* instruction only had an average pretest score of 629.18 (SD = 119.50), which was significantly higher (t (205) = 6.60, $p < .001$).

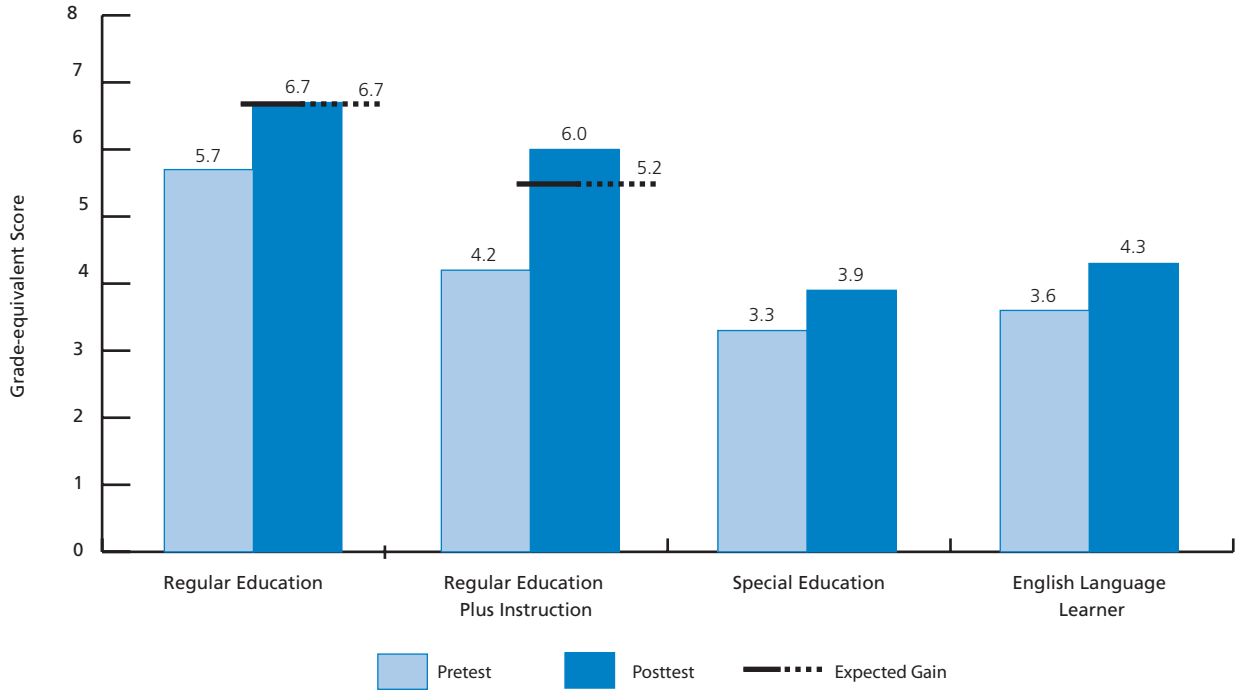
Because the groups differed significantly on their pretest scores, and therefore not equal in ability at the start of the intervention, an ANCOVA was conducted using the pretest score as the covariate to determine if adding classroom instruction to the *Learning 100* instruction had a significant impact on posttest scores. The analysis found that there was a significant effect of the covariate indicating that students' pretest scores were predictive of their posttest score (F(1,207) = 122.86, $p < .001$). However, once the impact of the pretest score had been removed, there was not a significant difference between the two groups on the STAR Assessment posttest, indicating that the addition of the classroom instruction did not impact achievement significantly. However, in practical terms, the students who received the additional classroom instruction experienced more growth throughout the year than with *Learning 100* alone, gaining almost two grade levels. This was extra beneficial to this group, given they had begun the year at a significantly lower pretest level and increased their posttest scores enough to statistically close the achievement gap. Therefore, the added educational benefit of the classroom instruction cannot be entirely dismissed. Furthermore, this analysis indicates that for regular education students reading at the lowest levels, *Learning 100* plus classroom instruction may benefit the lowest readers the most.

Kentucky

A multiyear evaluation of the *Learning 100* program was conducted at a high school in an urban city in Jefferson County, Kentucky, to determine the impact of *Learning 100* on the reading achievement scores of struggling readers. The initial evaluation took place during the 2000-2001 school year with ninth-grade students who were enrolled in a 12-week reading improvement program. In order to accommodate all of the students, three 12-week sessions of *Learning 100* were conducted. There were 433 ninth-grade students enrolled during the 2000-2001 school year. The demographics for the

Figure 3

Actual versus expected grade-level gain by classification on the STAR Assessment after using *Learning 100*



ninth grade were approximately as follows: 69% Caucasian, 28% African American, 3% Asian, less than 1% Native American, and 0% Hispanic. Total enrollment in the high school was 1,476 during the 2000-2001 school year and the student population had the following demographics: 73% Caucasian, 24% African American, 1% Asian, and less than 1% Hispanic, and Native American. The school had a free- and reduced-lunch rate of 37%.

During each 12-week reading improvement course, students used the *Learning 100* program independently for half of the class and spent the second half of the class working with a reading instructor. The quasi-experimental design of the evaluation included a pre-and posttest assessment of the students' reading abilities using the Test of Adult Basic Education (TABE), which was administered at the beginning and end of the 12-week period. Data were collected on reading comprehension ability and vocabulary knowledge.

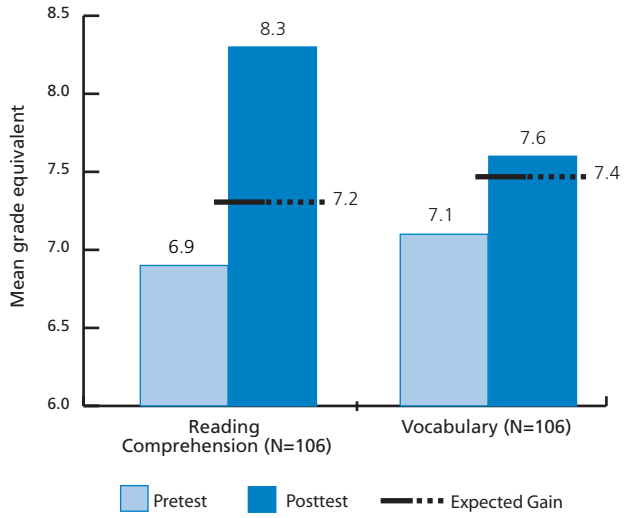
Given there was wide variability in the academic diversity of the student population participating in the program, some students achieved the highest level on the pretest TABE. Because this left no room for improvement, their scores were removed from the analysis. Additionally, students whose scores had fluctuated dramatically from pre- to posttest (increases of five grade levels and decreases of two grade levels) were removed from the data analysis due to the students possibly not putting forth the required effort. This left 106 students for the final analyses.

For a 12-week instructional period, it is expected that students' reading skills would naturally improve 0.3 grade levels without the aid of a specialized reading course. Students who participated in the reading improvement course had a reading comprehension grade level median of 6.9 at the beginning of the course and excelled to reach a median score of 8.3 by the end of the 12 weeks. This is a considerably larger gain[9] than would be expected by normal standards, which would predict a gain only to 7.2. The

[9] No statistical analyses were conducted on this data, all reported analyses were obtained from personnel at the participating high school. The original data was not available at this time, so only generalizations can be made about the significance of these gains.

vocabulary score of these students improved from 7.1 to 7.6, a gain of 0.5 grade levels. Although this level of improvement is not as large, it is greater than would be expected during a 12-week period. See Figure 4 for a graphical representation of this data.

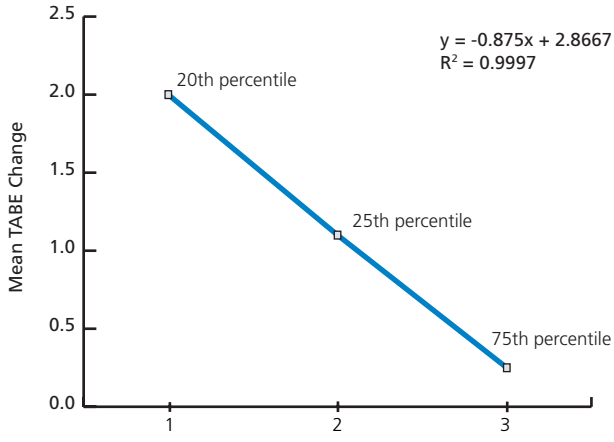
Figure 4
Actual versus expected gains based on TABE scores



Because of the range of academic ability included in the sample, the students' TABE gain scores were also broken down into subgroups by percentile rank. Overall, the results demonstrated that students at the lower percentile rank benefited more from the *Learning 100* program than students at the higher percentile ranks. Students at the 20th percentile improved 2.0 grade levels (from 3.75 to 5.75), students at the 25th percentile improved 1.1 grade levels (4.9 to 6.0), while students at the 75th percentile only improved .25 grades (10.65 to 10.9), which is less than would be normally expected. As Figure 5 displays, there was almost a perfect linear relationship between the students' percentile rank and their gain on the TABE. The more the students needed the help the more they benefited from the program.

For the vocabulary scores, students improved 0.5, 0.5, and 1.0 grade levels at the 25th, 50th, and 75th percentile levels respectively. Although not as impressive as the growth seen on the reading comprehension subtest, the level of growth was

Figure 5
Amount of change in TABE reading comprehension scores by percentile rank



consistently better at each percentile ranking than would be expected in 12 weeks without specialized instruction. In this case, the *Learning 100* program was better at increasing vocabulary knowledge for higher-achieving students.

The second evaluation of the *Learning 100* program at this school was conducted during the 2003-2004 school year. Data from the fall and spring were analyzed separately.[10] There were 469 ninth-grade students enrolled during the 2003-2004 school year. The demographic breakdown of the ninth-grade class was as follows: 71% Caucasian, 24% African American, 3% Hispanic, 1% Asian, and less than 1% Native American. Total enrollment in the high school was 1,356 during the 2003-2004 school year with 73% Caucasian, 23% African American, 2% Hispanic, 2% Asian, and less than 1% Native American. No estimate of the free- and reduced-lunch rate could be obtained at this time. The same implementation procedure was used as in the first evaluation in which students participated in a 12-week course using *Learning 100*. Students were administered the TABE at the beginning and end of the 12-week session. Data were collected on reading comprehension ability, vocabulary knowledge, and a total composite score during the 12-week period.

[10] No statistical analyses were conducted on this data because only grade equivalent scores were available at this time. Therefore only generalizations can be made about the significance of these gains.

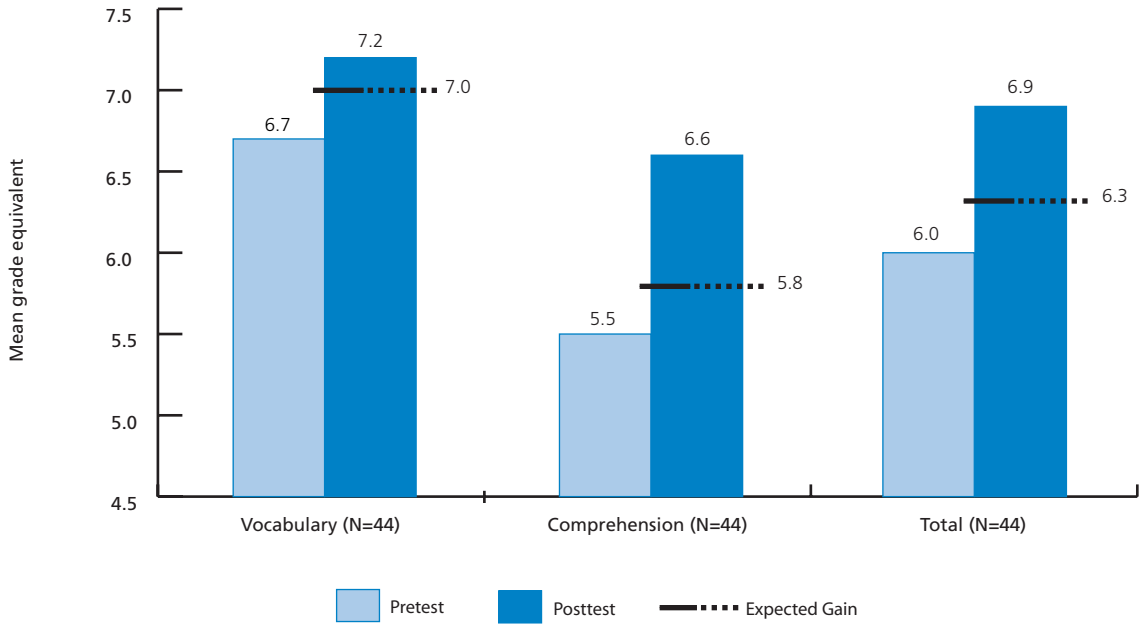
Because of the likelihood of academic diversity in the population, the same type of data restriction protocols were employed during the data analysis as in the first evaluation resulting in student data removal from the sample.

Data from 44 students who participated in the *Learning 100* program during the fall semester of the 2003-2004 school year were examined. The expected increase in grade level on the TABE for all subtests was .3 for the 12-week period. Students who participated in the *Learning 100* program scored on average a 6.7 grade-equivalent score on the test of vocabulary knowledge at the pretest and rose to 7.25 by the posttest. This gain surpassed the expected gain of 7.0, which would be expected over a 12-week implementation period. Students' grade-equivalent comprehension scores rose from an average of 5.5 to 6.6, which greatly exceeded the expected .3 gain, to 5.8. The students' total composite score increased on average .9 grade levels, three times the expected rate, from 6.05 to 6.9. See Figure 6 for these results. For this group of students, the *Learning 100* program had a greater impact on comprehension ability than vocabulary knowledge. Despite the lesser gains seen in vocabulary improvement, the average gain was higher than the expected gain of .3 grade levels, indicating on average students experienced growth in vocabulary.

Sixty students participated in the spring semester evaluation of *Learning 100* during the 2003-2004 school year. As with the previous evaluations, the 12-week period would be expected to yield .3 grade levels of gain on the TABE subtests. After using the *Learning 100* program for the 12-week implementation period, the students' vocabulary knowledge increased on average .6 grade levels, twice the expected amount, rising from a grade-equivalent score average of 7.6 to 8.2. Comprehension ability similarly rose on average from a grade-equivalent score of 7.9 to 8.4. Although the effect on comprehension ability was less than that on vocabulary, the gains made by the students were still greater than would be expected through normal growth. Students experienced gains in the total composite score as well. The average grade-equivalent score for the total composite test increased from 7.7 to 8.2, which was greater on

Figure 6

Learning 100 actual versus expected gains based on TABE scores



average than would have been expected after 12 weeks of instruction without an intervention program. See Figure 7 for these results. The *Learning 100* program had equal effects on vocabulary knowledge and reading comprehension ability for this population of students. The differential effects seen in achievement growth with this population, as opposed to the past evaluations at the same school, may be due to the fact that this group of students began the evaluation with higher levels of achievement, which may have restricted the level of overall improvement.

Taken as a whole, the three evaluations from this district indicate that *Learning 100* is an effective program for increasing ninth-grade student reading achievement for struggling readers above and beyond what would normally be expected. Although it was not consistent throughout all three evaluations, the *Learning 100* program seemed to have a larger impact on reading comprehension ability than vocabulary. Moreover, as the first evaluation demonstrated, the program was most beneficial for the lowest-achieving students increasing their scores by two full grade levels.

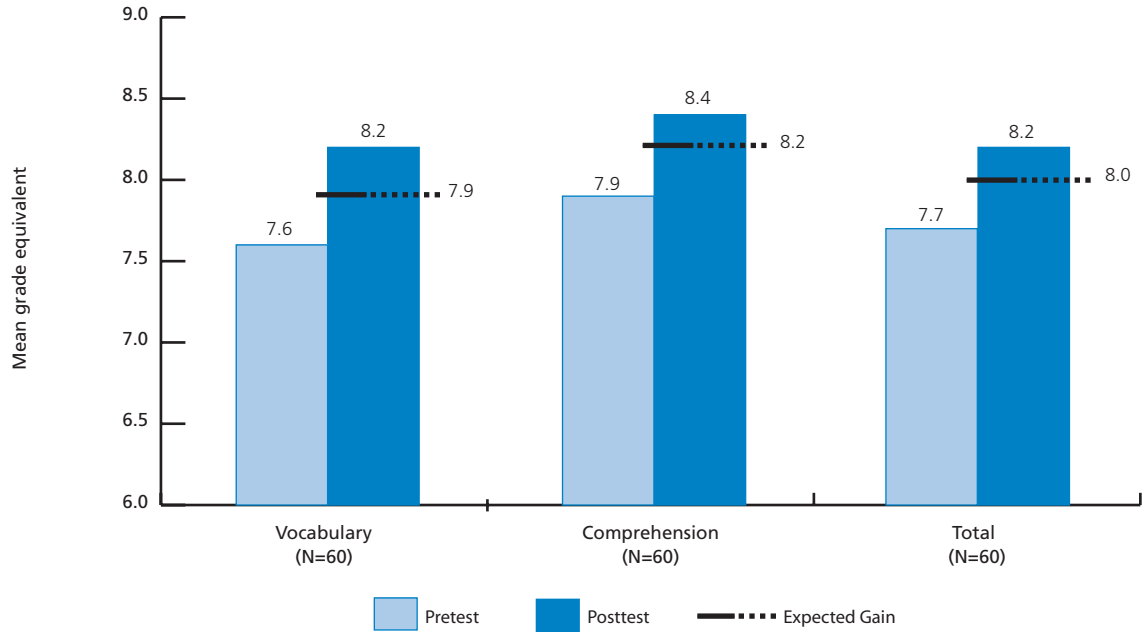
Louisiana

An efficacy study was conducted at the high school level in the Monroe City School District in urban northeast Louisiana during the 2002-2003 school year. Monroe City School District has three high schools. The *Learning 100 System* was purchased by the school district to support reading remediation classes across multiple grades in two high schools during the 2002-2003 school year. Although the initial plan for the study was to measure the impact of the program in both high schools, because of teacher illness, students in one school did not use the *Learning 100 System* and are not included in this analysis.

Forty students in grades 9 through 11 identified by their high school staff for reading remediation participated in the study, including 19 special education students. The total enrollment of the high school completing this study was 933. Fifty-eight percent of students were African American and 41% were Caucasian. Thirty-five percent of the student body participates in the free- or reduced-lunch program.

Figure 7

Learning 100 actual versus expected gains based on TABE scores



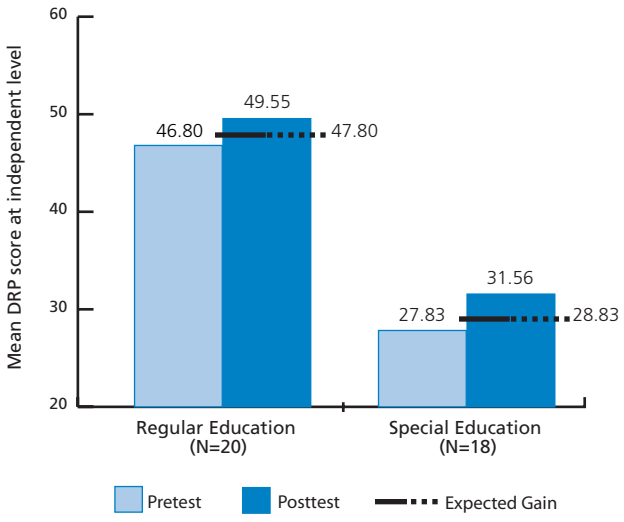
The study employed a quasi-experimental pretest/posttest design. The study was undertaken to determine if the *Learning 100 System* would be an effective tool with which to improve the reading achievement of high school students reading below grade level. Students' general reading achievement were evaluated using the Degrees of Reading Power (DRP) assessment as a pretest in October 2002 and a posttest in May 2003.

Overall, remedial reading students in this study utilized the *Learning 100 System* for nearly 32 hours over the course of the school year. Two of the 40 students were dropped from the study. One student was dropped because of incomplete data, while a second was eliminated from the sample, post-hoc, as an extreme statistical outlier. The Independent Level DRP Score (p = .9) was used in the analysis. This level defines reading independence as the ability to comprehend 90% of the material being read. Scoring using the DRP Independent Level is considered by the test publisher to be the most rigorous standard against which to evaluate DRP scores. The sample was split into regular education (n=20) and special education (n=18) for the purposes of analysis. A paired sample *t*-test was calculated to examine mean differences between pre- and posttest scores. Effect sizes were calculated using Cohen's *d*[11].

The mean DRP score for Regular Education students averaged 46.80 (*SD* = 10.65) on the pretest and 49.55 (*SD* = 9.41) on the posttest. The increase in score was found to be significant (t (19) = 2.48, p = .02; effect size = .27). The mean DRP score for special education students increased from a pretest score of 27.83 (*SD* = 12.78) to a posttest score of 31.56 (*SD* = 14.65) and is considered to be highly significant (t (17) = 3.17, p =.005; effect size = .27). See Figure 8 for these results. DRP technical reports state that students are expected to improve by one to two DRP units in a typical high school environment. The students in this study who utilized the *Learning 100 System* improved their score by 3.21 on average, well above the expected score gain.

The analysis demonstrates that the *Learning 100 System* facilitated improvement in students' reading comprehension skills. Even students with a special education classification were able to benefit from the *Learning 100 System*. Furthermore, the *Learning 100 System* appears to have contributed to higher than expected reading growth for both regular and special education students in a high school remedial setting.

Figure 8
Actual versus expected gains based on Degrees of Reading Power assessment after using *Learning 100* for one year



North Carolina

A 6-month evaluation of the *Learning 100* program took place in four high schools in a county school district located in the south-central area of rural North Carolina during the 2004–2005 school year. The district was demographically diverse. Forty-four percent of students were Caucasian, 34% of students were African American, 21% of students were Hispanic, and fewer than 1% each of students were Asian and American Indian[12]. Ninth-grade students who did not pass the eighth-grade end-of-grade reading test were required to participate in at least one semester[13] of a *Learning 100* class during the ninth grade as a separate course. Ninety-four

[11] Refer to Footnote 3 for the practical evaluation of a Cohen's *d* effect size.

[12] Demographics were gathered through the National Center for Education Statistics, based on 2003–2004 preliminary data.

[13] After the initial semester, all but four students volunteered to continue the class into the next semester.

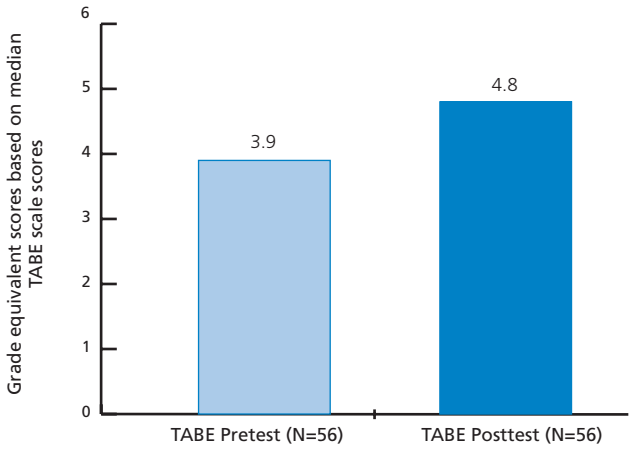
students qualified for and participated in this study. No additional systematic or formal remedial reading curriculum was in place during the implementation period. Teachers could aid students as needed, but the *Learning 100* program was the only source of classroom instruction used. After a teacher-orientation and training session, all participating students were pretested with the *Test of Adult Basic Education* (TABE; CTB/McGraw-Hill, 2003) (Complete Battery Level D, Form 9) and *Degrees of Reading Power* (DRP, Form J4) and began using the program in October of 2004. Students spent approximately 45 minutes engaged with the program on a daily basis. After 6 months, students were posttested with the TABE (Complete Battery Level D, Form 10) and DRP (Form K4).

Test of Adult Basic Education Results

Only the 61 students who had both a TABE pretest and posttest score were included in the dataset for further analysis. This created a sample of 61 students. The students' numbers correct on the pretest and posttest were converted into scale scores for further comparison. Because of apparent fluctuations in the data outside the bounds of typical variation, which could be due to a variety of factors beyond normal control, the data were initially organized by the amount of change from pretest to posttest. Because a student's true score can range from a standard error of measurement (*SEM*) below or above what is reported, an outlier in the data was defined as a posttest score outside the bounds of two average *SEMs* below or above the *SEM* range of the pretest or posttest score. The average *SEM* range in the data was 63 points below or above the average student score.

For example, a student scoring a pretest scale score of 434 would have a true score in the range of 407 to 461, which is a *SEM* of 27 below and above the reported scale score[14]. The group average *SEM* of 63 points below the lowest end of the range of scores for this student would be 344, 63 points below 407. If a student's posttest score was below 344, then the change from pretest to posttest would be considered well below what would have been thought of as a reasonable decrease, based on the estimation of the student's true score and the average *SEM* in this population. The same justification was used for scores increasing at a rate of change that would be considered uncharacteristic for this dataset. Within this dataset, five such outliers were identified and removed before the data were further analyzed, resulting in a final sample of 56 students.

Figure 9
Comparison between pre- and posttest TABE grade equivalent scores for ninth-grade struggling readers after using *Learning 100* for six months



Because potentially extreme scores were still present in the dataset after the removal of calculated outliers, the median was used to define the central tendency of change in the data from pretest to posttest in the dataset. Students had a median of 460 (*SD* = 80.33) on the pretest and 484 (*SD* = 60.26) on the posttest. These scale-score medians converted into grade-equivalent scores of 3.9 and 4.8, respectively (see Figure 9).

In a typically developing population, reading scores would be expected to increase by an average of 1.0 grade levels during a 9-month period, with no outside intervention. In a 6-month time span, an increase of .6 grade levels would therefore be expected. Because the students in this study were reading well below the expected grade level for ninth-grade students, a typical increase of .6 grade levels would not likely be expected within this population. However, with the aid of *Learning 100*,

[14] *SEM* obtained from the TABE Forms 9 and 10, Norms book, for the Complete Battery and Survey, All levels published by CTB/McGraw-Hill.

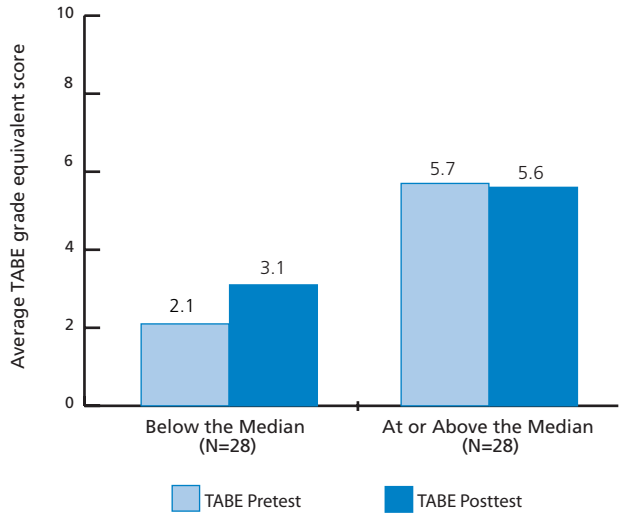
these students were able to increase their grade level score by .9 points, almost a full grade-level increase after only 6 months of implementation.

Because inferential statistics are calculated using average scores, the mean pretest and posttest scores were also calculated for this dataset. Students had an average pretest score of 448 (*SD* = 80.33) and an average posttest score of 471 (*SD* = 60.26). A paired *t*-test evaluation found that the difference between these scores was not significant. However, if the *t*-test results are examined (t (110) = 1.70, p = .09, *ES* = .32), it can be seen that there is a trend toward significance and an effect size that is above the educationally significant level. Additionally, the practical conversion to grade-equivalent scores indicated that, for this population of below-average readers, there were impressive gains made after using *Learning 100* for only 6 months.

Due to the variability in the reading ability of the students in the dataset, it was hypothesized that the *Learning 100* program may have had a differential effect, depending on a student's reading ability at the beginning of the implementation. Students were divided into two groups: those whose score fell below the median pretest score and those whose score fell at or above the median pretest score. Students whose pretest score fell below the median, students with the lowest reading ability, had an average pretest score of 385 (*SD* = 62.61) and a posttest score of 433 (*SD* = 56.66). The difference between these scores was both highly statistically significant and meaningfully significant (t (54) = 2.97, $p < .01$, *ES* = .80). For students with a higher reading ability, students whose pretest scores fell at or above the median, the average pretest score was 511 (*SD* = 31.51) and the average posttest score was 509 (*SD* = 37.78). The difference between the pretest and posttest scores was not significant and the slight decrease in score was not considered meaningful.

Turning to grade-equivalent scores (see Figure 10), students' whose pretest scores fell below the median had a beginning grade-equivalent score of 2.1 and ended the 6-month implementation with a grade-equivalent score of 3.1—a gain of 1.0 grade levels. Students who had a pretest score at or above

Figure 10
Comparison between pre- and posttest TABE grade equivalent scores for ninth-grade struggling readers after using *Learning 100* for six months



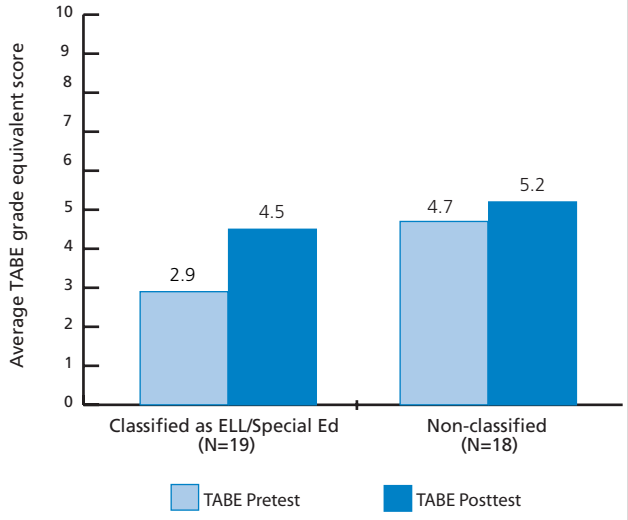
the median decreased .1 grade-equivalent scores, moving from 5.7 to 5.6; however, this would not be considered a meaningful drop. Taken together, these scores indicate that the students at the lowest levels of reading ability benefited much more from the *Learning 100* program than the students who were relatively better readers from the start of the implementation period. The students at the lowest level of reading ability were able to make gains of an entire year in 6 months' time, which, as stated previously, is not typical for such below-average readers.

Because the population of students participating in the evaluation of the *Learning 100* program included a high proportion of students labeled as English-language learnersw (ELL) or in need of special education services, the scores from these students were isolated further from the dataset to determine if the program differentially effected these subpopulations. Thirty-seven of the students had definite classifications reported, whereas 19 were reported as N/A. Of the 37 students with reported classifications, 19 were considered ELL or in need of special education. The remaining 18 were reported as nonclassified (i.e., neither ELL nor Special Ed).

The students classified as neither ELL nor Special Education had an average pretest score of 480 (*SD* = 86.62) and a posttest score of 494 (*SD* = 44.60), which translated into grade-equivalent scores of 4.7 and 5.2, respectively. The gain from pretest to posttest for this group was not significant. The students that were classified as either ELL or Special Education scored on average 427 (*SD* = 69.39) on the pretest and 474 (*SD* = 53.06) on the posttest, which was a significant gain (t (36) = 2.37, $p < .01$, *ES* = .76). This converted to a gain of from 2.9 to 4.5 on a grade-equivalent scale, which illustrates a gain of one and a half grade levels over the 6-month implementation period. See Figure 11 for a comparison between students who were classified as ELL/Special Ed and nonclassified students (i.e., neither ELL nor Special Ed).

Figure 11

Comparison between pre- and posttest TABE grade equivalent scores for ninth-grade struggling readers after using *Learning 100* for six months: ELL and special education students versus non-classified students



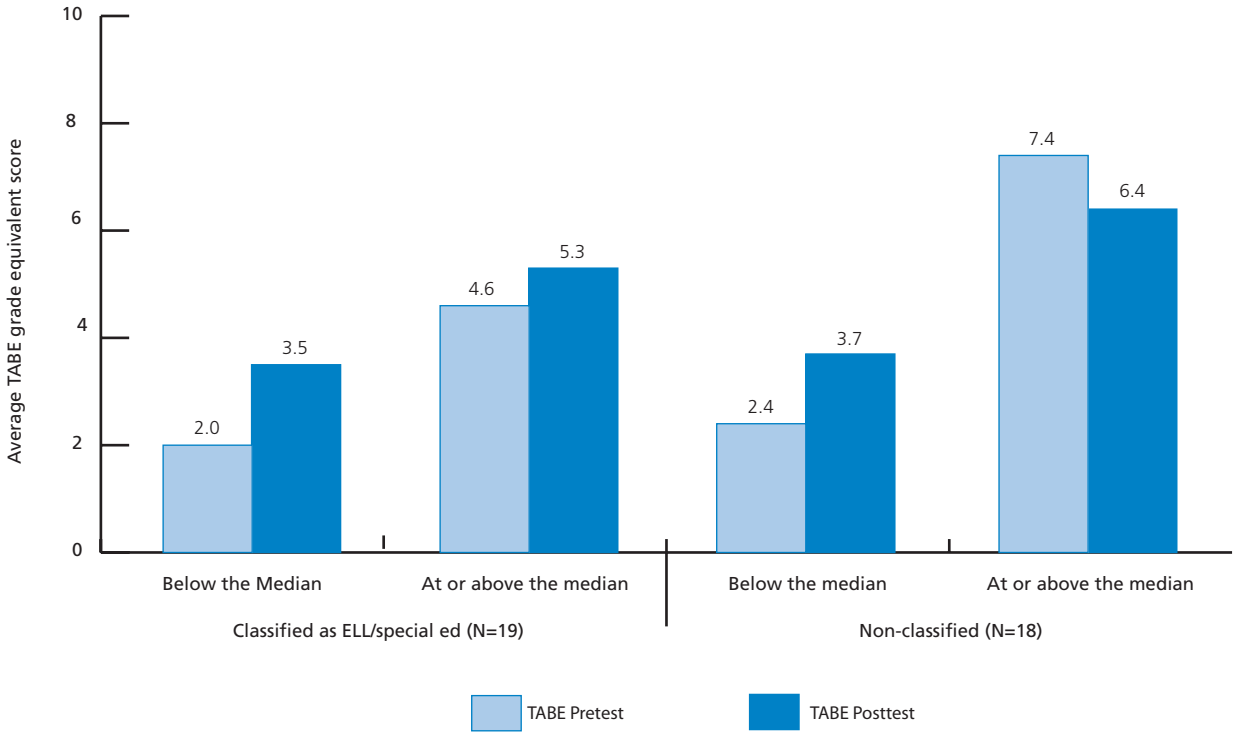
The graphs in Figure 11 show that the *Learning 100* program had a differential impact on students that depended on their classification and that benefited students who were in need of special services more than the general-education students who grew by only .5 points. Despite the lower level of growth, a gain of .5 in this population would still be regarded as a considerable amount of gain for struggling readers.

Because the *Learning 100* program had shown to be more effective with the lowest achieving students in the earlier analyses, the classification subpopulations were examined for this effect as well. Within the nonclassified group (i.e., neither ELL nor Special Ed), students reading below the median pretest score had a pretest score of 406 (*SD* = 79.19) and a posttest score of 455 (*SD* = 33.80), which was not a significant gain from pretest to posttest. Despite the lack of significance, which may be due to the smaller sample size, the gain in grade equivalency was impressive, rising from 2.4 to 3.7—a gain of 1.3 years. This was not the case for the higher achieving students who read at or above the median. They began the implementation with an average pretest score of 539 (*SD* = 22.56), which on average dropped to 526 (*SD* = 19.91). This loss was not significant and may be due to motivational factors.

For the students classified as either ELLs or Special Education who were reading below the median on the pretest, the average pretest score was 368 (*SD* = 52.55), which rose significantly to the average posttest score of 445 (*SD* = 54.31), t (16) = 3.04, $p < .01$, *ES* = 1.44. This translated into a grade equivalent score gain from 2.0 to 3.5. Not only was the grade-equivalent score gain large (1.5 years), the gain was also statistically meaningful, resulting in a large effect size. Even though these students were considered to be in need of special services, students who were reading at or above the median on the pretest in this group did not experience a significant gain from pretest to posttest, moving from an average pretest score of 480 (*SD* = 25.69) to an average posttest score of 501 (*SD* = 36.84). This gain represented a grade-level gain of .7, which although large for this population of struggling readers, was overall less than the gain for the lowest achieving readers. Taken together, the same pattern emerges as before, which indicates that the *Learning 100* program is benefiting the lowest achieving students in this population despite their classification. If classification is taken into consideration, the ELLs and Special Education students are benefiting the most, especially if they are reading below the median reading level in the population (see Figure 12).

Figure 12

Comparison between pre-and posttest TABE equivalent scores for ninth-grade struggling readers after using *Learning 100* for six months: ELL and special education students versus non-classified students reading below and above the median



Degrees of Reading Power Results

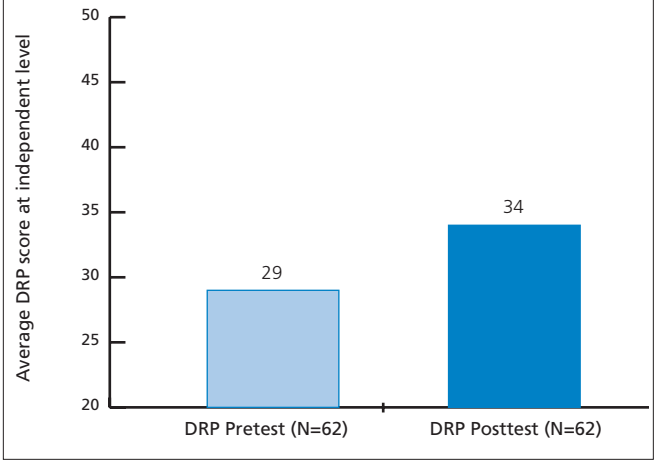
Data were compiled for 62[15] students who had both a pretest and posttest scores on the *Degrees of Reading Power* (Touchstone Applied Science Assoc., 2000). Although there was not a time limit given for the administration of the DRP, students were allowed to stop at the point of frustration. Because the DRP test gets progressively more difficult, many students stopped prior to completing the test. The number of attempted items on the pretest ranged from 7 to 70 (out of 70 possible questions) and 21 out of 70 on the posttest, indicating that students completed more questions by the end of the *Learning 100* evaluation. No clear outliers were identified in the dataset. The number correct for each student on the pretest and posttest were calculated and averaged[16]. Students scored on average 18.35 (*SD* = 11.99) attempted questions correct on the pretest and their score rose to 23.37 (*SD* = 10.15) questions correct on the posttest. These averages were transformed into DRP units at the Independent Level of Reading (P = .9) for further analysis[17]. This level defines reading independence as the ability to comprehend 90% of the material being read. Scoring using the DRP Independent Level is considered by the test publisher to be the most rigorous standard against which to evaluate DRP scores. Upon conversion, students had a score of 29 DRP units on the pretest and a score of 34 on the posttest (see Figure 13).

[15] Because the DRP and TABE tests were given at different times for the pretest and posttest administrations, the 62 students for whom there was complete DRP testing data do not necessarily overlap the 52 students that had complete TABE data.

[16] To aggregate DRP data to calculate statistical differences, the publisher of the DRP advises that such measures should be calculated on the raw number correct scores and then resulting means should be converted to DRP scores at the particular level of chosen comprehension ability. According to the publisher, "this procedure avoids cumulative errors due to rounding and/or possible ceiling and floor effects for a given DRP test form" (Touchstone Applied Science Assoc., 2000, p. 26)

[17] The Independent Reading Level, P = .9, was used for all DRP evaluations in this report.

Figure 13
Comparison between pre- and posttest DRP scores for ninth-grade struggling readers after using *Learning 100* for six months

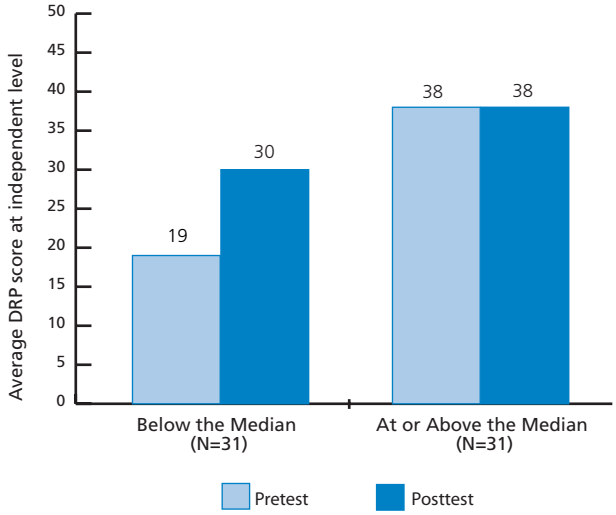


As stated in the *Degrees of Reading Power Handbook* (Touchstone Applied Science Assoc., 2000), in high school, about 1 to 2 DRP units of growth are expected for students between the fall and spring test administrations. Because the students participating in this study were reading at a below-average reading ability, the amount of growth expected without added support would likely be less than this. With the addition of supplemental help, the level of expected growth may vary, depending upon the individual differences of each student. However, because of the variation seen in the beginning reading ability of participating students within the dataset, the normal expectation of growth within a year for a high school student was used to judge the practical effects of the implementation of the *Learning 100* program. Students' scores rose on average by five DRP units within the course of the 6-month implementation of the *Learning 100* program, which is beyond the expected growth of two DRP units. A paired sample *t*-test with an alpha level set at .05 was used to compare the students' pretest to posttest growth performance using the students' number correct of attempted items. It was found that the growth from pretest to posttest was significant ($t(61) = 3.62$, $p = .001$, $ES = .45$), suggesting that students made tangible gains overall in reading comprehension while using the *Learning 100* program.

To test the effect of differential impact of prior reading ability on DRP scores, students were divided into groups including those who fell below the median reading ability on the pretest and those who fell at or above the median reading ability. Students who read below the median reading level at pretest had a mean pretest DRP number correct score of 8.74 ($SD = 4.54$) and a posttest score of 18.45 ($SD = 7.81$). The growth from pretest to posttest was highly significant ($t(30) = 6.35$, $p < .001$, $ES = 1.52$). Students who read at or above the median reading score at the pretest had an average pretest DRP number correct score of 27.97 ($SD = 8.99$), which grew slightly to 28.29 ($SD = 9.91$); however, this growth was not statistically significant. Additionally, students reading at the lowest reading levels attempted more than twice as many items on the posttest than on the pretest (44% to 90%), scoring a higher percentage of those attempted items correct. Students who were reading at a higher level attempted the same percentage of items from pretest to posttest (82%). By the posttest evaluation, the students who were initially reading at the lower reading levels attempted as many items as those students who began the year reading more proficiently, reflecting a possible increase in motivation levels of the lower ability readers over the course the implementation period.

Translating the average number correct scores into DRP units (see Figure 14), students whose

Figure 14
Comparison between pre- and posttest DRP scores above and below the median for ninth-grade struggling readers after using *Learning 100* for six months



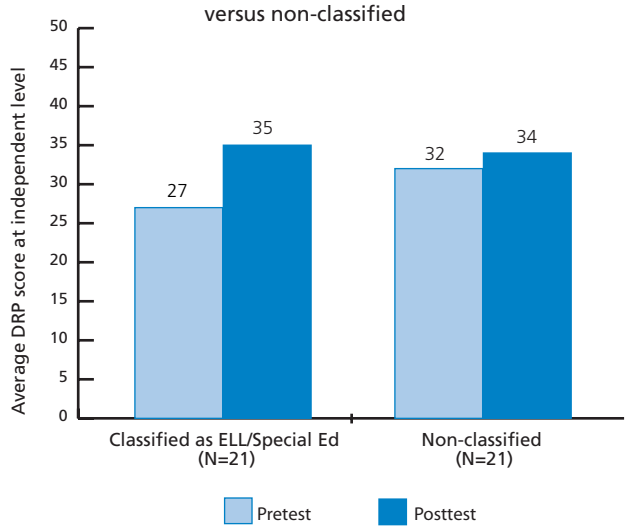
pretest score fell below the median pretest reading ability had a beginning DRP unit score of 19. The score rose to 30, which was a much higher gain than the expected 2 DRP units for a high school student. Students classified as reading at or above the median reading level at the pretest experienced no change in DRP unit from pretest to posttest, scoring 38 at both assessment times. As seen with the TABE analysis previously discussed, these data suggest that students reading at the lowest reading levels benefited more from the *Learning 100* program than did students reading at higher levels. They made more gains in reading comprehension ability within the 4-month evaluation time and experienced a possible change in motivation and confidence in their own reading ability.

To determine if there was a differential effect of the *Learning 100* program for students with special needs, scores for individuals who were classified as either ELLs or needing special education services were analyzed and compared to students who were not classified with these needs. Of the students with complete DRP test scores, 21 students were classified as either ELLs or needing special education services, and 21 students were nonclassified (i.e., neither ELLs nor Special Ed). Twenty students received a classification of N/A.

Students who received a classification of either ELL or needing special education services had an average number correct on the DRP pretest of 15.77 (*SD* = 10.6), which rose to an average number correct of 23.86 (*SD* = 11.14) on the posttest. This gain from pretest to posttest was significant ($t(20) = 3.42$, $p = .002$, $ES = .76$). Students who were not classified as needing special services scored an average of 21.29 (*SD* = 11.17) questions correct on the pretest DRP, which rose to a score of 24.33 (*SD* = 8.57) on the posttest. The difference between pretest and posttest administrations for nonclassified students (i.e., neither ELL nor Special Ed) was not significant.

Converting these scores into DRP units, classified students rose from a score of 27 on the pretest to a score of 35 on the posttest. The gain for ELLs and students in need of special education services after 6 months of *Learning 100* instruction is larger than the expected gain of 2 DRP units. Students who were nonclassified (i.e., neither ELL nor Special Ed) began the year with a score of 32, which rose to 34 after 6 months. Although this gain is not statistically significant, it is within the expected yearly gains for a high school student. As can be seen in Figure 15, *Learning 100* had a more substantial effect for ELLs and students needing special education services than for nonclassified (i.e., neither ELL nor Special Ed) students. Although nonclassified students made expected gains, the classified students made gains far above expectation.

Figure 15
Comparison between pre- and posttest DRP scores for ninth-grade struggling readers after using *Learning 100* for six months: ELL and special education students versus non-classified



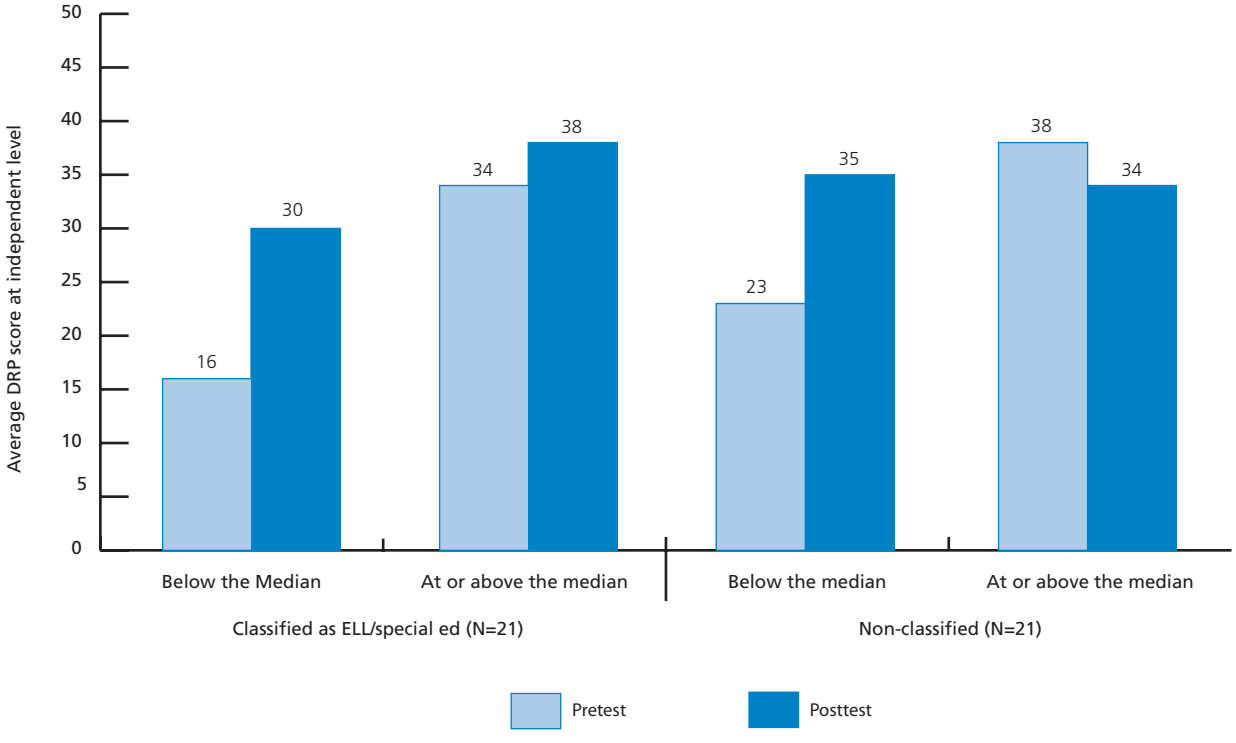
To determine if beginning reading ability was a factor within these classifications, students were further divided into groups including students reading below the median reading level at the pretest and students reading at or above the median. Classified students reading below the median had an average score of 7.20 (*SD* = 4.47) questions correct on the DRP pretest, which rose to a score of 19.00 (*SD* = 8.72), which was a significant gain ($t(9) = 4.33$, $p = .002$, $ES = 1.70$). Classified students who were reading at or above the median pretest reading ability scored, on average, 23.55 (*SD* = 6.67) questions correct on the pretest and 28.27 (*SD* = 11.60) questions correct on the posttest. This difference was not significant. Students who were nonclassified as either ELLs or needing special

education services and who were reading below the median pretest level for these students scored on average 12.30 (*SD* = 5.68) questions correct on the DRP pretest, which rose significantly to 24.60 on the posttest ($t(9) = 5.44$, $p < .001$). Nonclassified students reading at or above the median reading level scored 29.45 questions correct on the DRP pretest, and their scores dropped slightly to 24.09 on the posttest. This drop was significant ($t(10) = 2.24$, $p = .05$, $ES = .64$), but it just reached significance. However, it should be noted that these students attempted fewer questions on the posttest than on the pretest, indicating a possible decline in motivational factor as the cause of the significant decline from pretest to posttest. Both classified and nonclassified students reading at the lowest levels attempted a higher percentage of questions on the posttest than on the pretest; the classified students gained the most over time, with scores rising from 37% to 93% of questions attempted. In both cases, students scored more attempted questions correct on the posttest. This indicates that the lower reading level students possibly felt more confident in their reading ability and attempted to answer more questions after the 6-month evaluation of the *Learning 100* program.

As can be seen in Figure 16, when the average number correct scores are translated into DRP units, scores from classified students reading below the median pretest reading level rose from 16 to 30 DRP units over the course of 6 months. This was well above the expected increase during this time period. Scores from classified students reading at or above the median pretest reading ability rose as well but by a smaller margin, from 34 to 38. Although classified students' overall gain was smaller, it was still higher than expected. Nonclassified (i.e., neither ELLs nor Special Ed) students reading below the median pretest level began at a higher level of reading ability than students who were classified as either ELLs or needing special services, but their reading ability was still well below the typical reading level. These students experienced a large growth over time in terms of DRP units as well, rising from 23

Figure 16

Comparison between pre-and posttest DRP scores for ninth-grade struggling readers after using Learning 100 for six months: ELL and special education students versus non-classified students reading below and above the median



to 35, but the magnitude of the gain was not as large as that of the classified students who had began the program reading at a lower ability in general. Nonclassified students who were reading at or above the median reading level experienced a drop in DRP units, but this may be do to a decrease in motivation. Taken together, the same pattern emerges as from the TABE data analysis, which is that the *Learning 100* program seems to be differentially benefiting the students with the highest need and the lowest reading ability within the population.

Upon examining the complete TABE and DRP analyses together for this population, a clear pattern emerges suggesting that *Learning 100* was successful overall at contributing to growth in student reading achievement over time. There are also indications that *Learning 100* has a differential impact on student reading achievement, depending upon the prior reading ability of a student and the student's need for special services. *Learning 100* is more beneficial for students who began the year reading below the median reading ability of the class and for those students who were classified as either ELLs or in need of special education. Overall, this evaluation demonstrates that the *Learning 100* program has a significant and practical positive effect on the reading achievement of struggling readers and possibly contributes to the motivational factors involved in reading comprehension.

Bibliography

CTB/McGraw-Hill. (2003). *Test of adult basic education*. Monterey, CA: Author.

Slavin, R.E. (1986). Best-evidence synthesis: An alternative to meta-analysis and traditional reviews. *Educational Researcher, 15,* 5-11.

Touchstone Applied Science Assoc. (2000). *Degrees of reading power handbook*. Brewster, NY: Author.

This page intentionally left blank.

10M/NY/SPEC/11-05
9994113402 IND
9994113410 10PK

Harcourt Achieve
Rigby • Saxon • Steck-Vaughn

www.HarcourtAchieve.com
1.800.531.5015